



SEQUENCE LISTING

(1) GENERAL INFORMATION:

5 (i) APPLICANT: Eaton, Dan L.
de Sauvage, Frederic J.

(ii) TITLE OF INVENTION: HUMAN MPL LIGAND

10 (iii) NUMBER OF SEQUENCES: 30

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Genentech, Inc.
(B) STREET: 460 Point San Bruno Blvd
(C) CITY: South San Francisco
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 94080

20 (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: WinPatin (Genentech)

25 (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 08/223263
(B) FILING DATE: 04-Apr-1994
(C) CLASSIFICATION:

30 (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/196689
(B) FILING DATE: 15-FEB-1994

35 (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/185607
(B) FILING DATE: 21-JAN-1994

40 (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/176553
(B) FILING DATE: 3-JAN-1994

45 (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Winter, Daryl B.
(B) REGISTRATION NUMBER: 32,637
(C) REFERENCE/DOCKET NUMBER: P0871P3

50 (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 415/225-1249
(B) TELEFAX: 415/952-9881
(C) TELEX: 910/371-7168

(2) INFORMATION FOR SEQ ID NO:1:

5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 353 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

10 Met Glu Leu Thr Glu Leu Leu Leu Val Val Met Leu Leu Leu Thr
-21 -20 -15 -10

15 Ala Arg Leu Thr Leu Ser Ser Pro Ala Pro Pro Ala Cys Asp Leu
-5 1 5

20 Arg Val Leu Ser Lys Leu Leu Arg Asp Ser His Val Leu His Ser
10 15 20

25 Arg Leu Ser Gln Cys Pro Glu Val His Pro Leu Pro Thr Pro Val
25 30 35

30 Leu Leu Pro Ala Val Asp Phe Ser Leu Gly Glu Trp Lys Thr Gln
40 45 50

35 Met Glu Glu Thr Lys Ala Gln Asp Ile Leu Gly Ala Val Thr Leu
55 60 65

40 Leu Leu Glu Gly Val Met Ala Ala Arg Gly Gln Leu Gly Pro Thr
70 75 80

45 Cys Leu Ser Ser Leu Leu Gly Gln Leu Ser Gly Gln Val Arg Leu
85 90 95

50 Leu Leu Gly Ala Leu Gln Ser Leu Leu Gly Thr Gln Leu Pro Pro
100 105 110

55 Gln Gly Arg Thr Thr Ala His Lys Asp Pro Asn Ala Ile Phe Leu
115 120 125

60 Ser Phe Gln His Leu Leu Arg Gly Lys Val Arg Phe Leu Met Leu
130 135 140

65 Val Gly Gly Ser Thr Leu Cys Val Arg Arg Ala Pro Pro Thr Thr
145 150 155

70 Ala Val Pro Ser Arg Thr Ser Leu Val Leu Thr Leu Asn Glu Leu
160 165 170

75 Pro Asn Arg Thr Ser Gly Leu Leu Glu Thr Asn Phe Thr Ala Ser
175 180 185

Ala Arg Thr Thr Gly Ser Gly Leu Leu Lys Trp Gln Gln Gly Phe
190 195 200

Arg Ala Lys Ile Pro Gly Leu Leu Asn Gln Thr Ser Arg Ser Leu
5 205 210 215

Asp Gln Ile Pro Gly Tyr Leu Asn Arg Ile His Glu Leu Leu Asn
220 225 230

Gly Thr Arg Gly Leu Phe Pro Gly Pro Ser Arg Arg Thr Leu Gly
10 235 240 245

Ala Pro Asp Ile Ser Ser Gly Thr Ser Asp Thr Gly Ser Leu Pro
15 250 255 260

Pro Asn Leu Gln Pro Gly Tyr Ser Pro Ser Pro Thr His Pro Pro
265 270 275

Thr Gly Gln Tyr Thr Leu Phe Pro Leu Pro Pro Thr Leu Pro Thr
20 280 285 290

Pro Val Val Gln Leu His Pro Leu Leu Pro Asp Pro Ser Ala Pro
295 300 305

Thr Pro Thr Pro Thr Ser Pro Leu Leu Asn Thr Ser Tyr Thr His
25 310 315 320

C1
cont.

Ser Gln Asn Leu Ser Gln Glu Gly
325 330 332

30 (2) INFORMATION FOR SEQ ID NO:2:

35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1795 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TCTTCCTTACCATCTGCTCC CCAGAGGGCT GCCTGCTGTG CACTTGGGTC 50
CTGGAGCCCT TCTCCACCCG GATAGATTCC TCACCCCTTGG CCCGCCTTTG 100
45 CCCCCACCCCTA CTCTGCCAG AAGTGCAAGA GCCTAAGCCG CCTCCATGGC 150
CCCCAGGAAGG ATTCAAGGGGA GAGGCCCAA ACAGGGAGCC ACGCCAGCCA 200
50 GACACCCCGG CCAGA ATG GAG CTG ACT GAA TTG CTC CTC 239
Met Glu Leu Thr Glu Leu Leu Leu
-21 -20 -15

CCC ACC ACA GCT GTC CCC AGC AGA ACC TCT CTA GTC CTC 785
 Pro Thr Thr Ala Val Pro Ser Arg Thr Ser Leu Val Leu
 160 165

5 ACA CTG AAC GAG CTC CCA AAC AGG ACT TCT GGA TTG TTG 824
 Thr Leu Asn Glu Leu Pro Asn Arg Thr Ser Gly Leu Leu
 170 175 180

10 GAG ACA AAC TTC ACT GCC TCA GCC AGA ACT ACT GGC TCT 863
 Glu Thr Asn Phe Thr Ala Ser Ala Arg Thr Thr Gly Ser
 185 190 195

15 GGG CTT CTG AAG TGG CAG CAG GGA TTC AGA GCC AAG ATT 902
 Gly Leu Leu Lys Trp Gln Gln Gly Phe Arg Ala Lys Ile
 200 205

20 CCT GGT CTG CTG AAC CAA ACC TCC AGG TCC CTG GAC CAA 941
 Pro Gly Leu Leu Asn Gln Thr Ser Arg Ser Leu Asp Gln
 210 215 220

25 ATC CCC GGA TAC CTG AAC AGG ATA CAC GAA CTC TTG AAT 980
 Ile Pro Gly Tyr Leu Asn Arg Ile His Glu Leu Leu Asn
 225 230

30 C1 cont
 GGA ACT CGT GGA CTC TTT CCT GGA CCC TCA CGC AGG ACC 1019
 Gly Thr Arg Gly Leu Phe Pro Gly Pro Ser Arg Arg Thr
 235 240 245

35 CTA GGA GCC CCG GAC ATT TCC TCA GGA ACA TCA GAC ACA 1058
 Leu Gly Ala Pro Asp Ile Ser Ser Gly Thr Ser Asp Thr
 250 255 260

40 GGC TCC CTG CCA CCC AAC CTC CAG CCT GGA TAT TCT CCT 1097
 Gly Ser Leu Pro Pro Asn Leu Gln Pro Gly Tyr Ser Pro
 265 270

45 TCC CCA ACC CAT CCT CCT ACT GGA CAG TAT ACG CTC TTC 1136
 Ser Pro Thr His Pro Pro Thr Gly Gln Tyr Thr Leu Phe
 275 280 285

50 CCT CTT CCA CCC ACC TTG CCC ACC CCT GTG GTC CAG CTC 1175
 Pro Leu Pro Pro Thr Leu Pro Thr Pro Val Val Gln Leu
 290 295

CAC CCC CTG CTT CCT GAC CCT TCT GCT CCA ACG CCC ACC 1214
 His Pro Leu Leu Pro Asp Pro Ser Ala Pro Thr Pro Thr
 300 305 310

CCT ACC AGC CCT CTT CTA AAC ACA TCC TAC ACC CAC TCC 1253
 Pro Thr Ser Pro Leu Leu Asn Thr Ser Tyr Thr His Ser
 315 320 325

CAG AAT CTG TCT CAG GAA GGG T AAGGT TCTCAGACAC 1290
Gln Asn Leu Ser Gln Glu Gly
330 332

5 TGCCGACATC AGCATTGTCT CATGTACAGC TCCCTTCCCT GCAGGGCGCC 1340

CCTGGGAGAC AACTGGACAA GATTCCTAC TTTCTCCTGA AACCCAAAGC 1390

10 CCTGGTAAAA GGGATACACA GGACTGAAAA GGGATCATT TTTCACTGTA 1440

CATTATAAAC CTTCAGAAGC TATTTTTA AGCTATCAGC AATACTCATC 1490

AGAGCAGCTA GCTCTTGTT CTATTTCTG CAGAAATTG CAACTCACTG 1540

15 ATTCTCTACA TGCTCTTTT CTGTGATAAC TCTGCAAAGG CCTGGGCTGG 1590

CCTGGCAGTT GAACAGAGGG AGAGACTAAC CTTGAGTCAG AAAACAGAGA 1640

20 AAGGGTAATT TCCTTGCTT CAAATTCAAG GCCTTCCAAC GCCCCCATCC 1690

CCTTTACTAT CATTCTCAGT GGGACTCTGA TCCCATATTC TTAACAGATC 1740

TTTACTCTTG AGAAATGAAT AAGCTTCTC TCAGAAAAAA AAAAAAAA 1790

25 AAAAAA 1795

C
cont.
(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 42 amino acids
(B) TYPE: Amino Acid
(C) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

35 Leu Leu Leu Val Val Met Leu Leu Leu Thr Ala Arg Leu Thr Leu
-16 -15 -10 -5

40 Ser Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val Leu Ser Lys
1 5 10

Leu Leu Arg Asp Ser His Val Leu His Ser Arg Leu
15 20 25 26

45 (2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 390 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GAATTCCCTGG AATACCAGCT GACAATGATT TCCTCCTCAT CTTTCAACCT 50

5 CACCTCTCCT CATCTAAGAA TTG CTC CTC GTG GTC ATG CTT 91
Leu Leu Leu Val Val Met Leu
-16 -15 -10

10 CTC CTA ACT GCA AGG CTA ACG CTG TCC AGC CCG GCT CCT 130
Leu Leu Thr Ala Arg Leu Thr Leu Ser Ser Pro Ala Pro
-5 1

15 CCT GCT TGT GAC CTC CGA GTC CTC AGT AAA CTG CTT CGT 169
Pro Ala Cys Asp Leu Arg Val Leu Ser Lys Leu Leu Arg
5 10 15

20 GAC TCC CAT GTC CTT CAC AGC AGA CTG GTGA GAACTCCCAA 210
Asp Ser His Val Leu His Ser Arg Leu
20 25 26

CATTATCCCC TTTATCCGCG TAACTGGTAA GACACCCATA CTCCCAGGAA 260

25 GACACCATCA CTTCCCTCTAA CTCCTTGACC CAATGACTAT TCTTCCCATA 310

TTGTCCCCAC CTACTGATCA CACTCTCTGA CAAGAATTAT TCTTCACAAT 360

ACAGCCCGCA TTTAAAAGCT CTCGTCTAGA 390

30 (2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 390 base pairs
(B) TYPE: Nucleic Acid
35 (C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

40 TCTAGACGAG AGCTTTAAA TGGGGCTGT ATTGTGAAGA ATAATTCTTG 50

TCAGAGAGTG TGATCAGTAG GTGGGGACAA TATGGGAAGA ATAGTCATTG 100

45 GGTCAAGGAG TTAGAGGAAG TGATGGTGTGTC TTCCCTGGGAG TATGGGTGTC 150

TTACCAGTTA CGCGGATAAAA GGGGATAATG TTGGGAGTTC TCACCAAGTCT 200

GCTGTGAAGG ACATGGGAGT CACGAAGCAG TTTACTGAGG ACTCGGAGGT 250

50 CACAAGCAGG AGGAGCCGGG CTGGACAGCG TTAGCCTTGC AGTTAGGAGA 300

AGCATGACCA CGAGGAGCAA TTCTTAGATG AGGAGAGGTG AGGTTGAAAG 350

ATGAGGAGGA AATCATTGTC AGCTGGTATT CCAGGAATTC 390

5 (2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 332 amino acids
(B) TYPE: Amino Acid
10 (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val Leu Ser Lys Leu
15 1 5 10 15

Leu Arg Asp Ser His Val Leu His Ser Arg Leu Ser Gln Cys Pro
20 20 25 30

20 Glu Val His Pro Leu Pro Thr Pro Val Leu Leu Pro Ala Val Asp
35 35 40 45

Phe Ser Leu Gly Glu Trp Lys Thr Gln Met Glu Glu Thr Lys Ala
25 50 55 60

C' Gln Asp Ile Leu Gly Ala Val Thr Leu Leu Leu Glu Gly Val Met
Cont. 65 70 75

30 Ala Ala Arg Gly Gln Leu Gly Pro Thr Cys Leu Ser Ser Leu Leu
80 85 90

Gly Gln Leu Ser Gly Gln Val Arg Leu Leu Leu Gly Ala Leu Gln
35 95 100 105

Ser Leu Leu Gly Thr Gln Leu Pro Pro Gln Gly Arg Thr Thr Ala
40 110 115 120

His Lys Asp Pro Asn Ala Ile Phe Leu Ser Phe Gln His Leu Leu
125 130 135

45 Arg Gly Lys Val Arg Phe Leu Met Leu Val Gly Gly Ser Thr Leu
140 145 150

Cys Val Arg Arg Ala Pro Pro Thr Thr Ala Val Pro Ser Arg Thr
155 160 165

Ser Leu Val Leu Thr Leu Asn Glu Leu Pro Asn Arg Thr Ser Gly
50 170 175 180

Leu Leu Glu Thr Asn Phe Thr Ala Ser Ala Arg Thr Thr Gly Ser
185 190 195

Gly Leu Leu Lys Trp Gln Gln Gly Phe Arg Ala Lys Ile Pro Gly
200 205 210

5 Leu Leu Asn Gln Thr Ser Arg Ser Leu Asp Gln Ile Pro Gly Tyr
215 220 225

Leu Asn Arg Ile His Glu Leu Leu Asn Gly Thr Arg Gly Leu Phe
230 235 240

10 Pro Gly Pro Ser Arg Arg Thr Leu Gly Ala Pro Asp Ile Ser Ser
245 250 255

Gly Thr Ser Asp Thr Gly Ser Leu Pro Pro Asn Leu Gln Pro Gly
260 265 270

15 Tyr Ser Pro Ser Pro Thr His Pro Pro Thr Gly Gln Tyr Thr Leu
275 280 285

20 Phe Pro Leu Pro Pro Thr Leu Pro Thr Pro Val Val Gln Leu His
290 295 300

Pro Leu Leu Pro Asp Pro Ser Ala Pro Thr Pro Thr Pro Thr Ser
305 310 315

25 Pro Leu Leu Asn Thr Ser Tyr Thr His Ser Gln Asn Leu Ser Gln
320 325 330

30 Glu Gly
332

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 166 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

40 Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr
1 5 10 15

Leu Leu Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala
20 25 30

45 Glu His Cys Ser Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys
35 40 45

50 Val Asn Phe Tyr Ala Trp Lys Arg Met Glu Val Gly Gln Gln Ala
50 55 60

Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser Glu Ala Val Leu
65 70 75

5 Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu Pro
80 85 90

Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu
95 100 105

10 Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser
110 115 120

15 Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala
125 130 135

Asp Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg
140 145 150

20 Gly Lys Leu Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp
155 160 165

Arg
166

25 (2) INFORMATION FOR SEQ ID NO:8:

*C'
cont.*

30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1443 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

35 GAGTCCTTGG CCCACCTCTC TCCCACCCGA CTCTGCCGAA AGAAGCACAG 50

AAGCTCAAGC CGCCTCCATG GCCCCAGGAA AGATTCAGGG GAGAGGCC 100

40 ATACAGGGAG CCACTTCACT TAGACACCCT GGCCAGA ATG GAG 143
Met Glu
-21 -20

45 CTG ACT GAT TTG CTC CTG GCG GCC ATG CTT CTT GCA GTG 182
Leu Thr Asp Leu Leu Ala Ala Met Leu Leu Ala Val
-15 -10

50 GCA AGA CTA ACT CTG TCC AGC CCC GTA GCT CCT GCC TGT 221
Ala Arg Leu Thr Leu Ser Ser Pro Val Ala Pro Ala Cys
-5 1 5

GAC CCC AGA CTC CTA AAT AAA CTG CTG CGT GAC TCC CAC 260
 Asp Pro Arg Leu Leu Asn Lys Leu Leu Arg Asp Ser His
 10 15 20

5 CTC CTT CAC AGC CGA CTG AGT CAG TGT CCC GAC GTC GAC 299
 Leu Leu His Ser Arg Leu Ser Gln Cys Pro Asp Val Asp
 25 30

10 CCT TTG TCT ATC CCT GTT CTG CTG CCT GCT GTG GAC TTT 338
 Pro Leu Ser Ile Pro Val Leu Leu Pro Ala Val Asp Phe
 35 40 45

15 AGC CTG GGA GAA TGG AAA ACC CAG ACG GAA CAG AGC AAG 377
 Ser Leu Gly Glu Trp Lys Thr Gln Thr Glu Gln Ser Lys
 50 55

20 GCA CAG GAC ATT CTA GGG GCA GTG TCC CTT CTA CTG GAG 416
 Ala Gln Asp Ile Leu Gly Ala Val Ser Leu Leu Leu Glu
 60 65 70

25 GGA GTG ATG GCA GCA CGA GGA CAG TTG GAA CCC TCC TGC 455
 Gly Val Met Ala Ala Arg Gly Gln Leu Glu Pro Ser Cys
 75 80 85

30 CTC TCA TCC CTC CTG GGA CAG CTT TCT GGG CAG GTT CGC 494
 Leu Ser Ser Leu Leu Gly Gln Leu Ser Gly Gln Val Arg
 90 95

C
 35 CTC CTC TTG GGG GCC CTG CAG GGC CTC CTA GGA ACC CAG 533
 Leu Leu Leu Gly Ala Leu Gln Gly Leu Leu Gly Thr Gln
 100 105 110

40 GGC AGG ACC ACA GCT CAC AAG GAC CCC AAT GCC CTC TTC 572
 Gly Arg Thr Thr Ala His Lys Asp Pro Asn Ala Leu Phe
 115 120

45 TTG AGC TTG CAA CAA CTG CTT CGG GGA AAG GTG CGC TTC 611
 Leu Ser Leu Gln Gln Leu Leu Arg Gly Lys Val Arg Phe
 125 130 135

50 CTG CTT CTG GTA GAA GGT CCC ACC CTC TGT GTC AGA CGG 650
 Leu Leu Leu Val Glu Gly Pro Thr Leu Cys Val Arg Arg
 140 145 150

55 ACC CTG CCA ACC ACA GCT GTC CCA AGC AGT ACT TCT CAA 689
 Thr Leu Pro Thr Thr Ala Val Pro Ser Ser Thr Ser Gln
 155 160

60 CTC CTC ACA CTA AAC AAG TTC CCA AAC AGG ACT TCT GGA 728
 Leu Leu Thr Leu Asn Lys Phe Pro Asn Arg Thr Ser Gly
 165 170 175

TTG TTG GAG ACG AAC TTC AGT GTC ACA GCC AGA ACT GCT 767
Leu Leu Glu Thr Asn Phe Ser Val Thr Ala Arg Thr Ala
180 185

5 GGC CCT GGA CTT CTG AGC AGG CTT CAG GGA TTC AGA GTC 806
Gly Pro Gly Leu Leu Ser Arg Leu Gln Gly Phe Arg Val
190 195 200

10 AAG ATT ACT CCT GGT CAG CTA AAT CAA ACC TCC AGG TCC 845
Lys Ile Thr Pro Gly Gln Leu Asn Gln Thr Ser Arg Ser
205 210 215

15 CCA GTC CAA ATC TCT GGA TAC CTG AAC AGG ACA CAC GGA 884
Pro Val Gln Ile Ser Gly Tyr Leu Asn Arg Thr His Gly
220 225

20 CCT GTG AAT GGA ACT CAT GGG CTC TTT GCT GGA ACC TCA 923
Pro Val Asn Gly Thr His Gly Leu Phe Ala Gly Thr Ser
230 235 240

25 CTT CAG ACC CTG GAA GCC TCA GAC ATC TCG CCC GGA GCT 962
Leu Gln Thr Leu Glu Ala Ser Asp Ile Ser Pro Gly Ala
245 250

30 CTT AAC AAA GGC TCC CTG GCA TTC AAC CTC CAG GGT GGA 1001
Phe Asn Lys Gly Ser Leu Ala Phe Asn Leu Gln Gly Gly
255 260 265

35 CTT CCT CCT TCT CCA AGC CTT GCT CCT GAT GGA CAC ACA 1040
Leu Pro Pro Ser Pro Ser Leu Ala Pro Asp Gly His Thr
270 275 280

40 CCC TTC CCT CCT TCA CCT GCC TTG CCC ACC ACC CAT GGA 1079
Pro Phe Pro Pro Ser Pro Ala Leu Pro Thr Thr His Gly
285 290

TCT CCA CCC CAG CTC CAC CCC CTG TTT CCT GAC CCT TCC 1118
Ser Pro Pro Gln Leu His Pro Leu Phe Pro Asp Pro Ser
295 300 305

45 ACC ACC ATG CCT AAC TCT ACC GCC CCT CAT CCA GTC ACA 1157
Thr Thr Met Pro Asn Ser Thr Ala Pro His Pro Val Thr
310 315

ATG TAC CCT CAT CCC AGG AAT TTG TCT CAG GAA ACA TAGCGC 1199
Met Tyr Pro His Pro Arg Asn Leu Ser Gln Glu Thr
320 325 330 331

50 G GGCAC TGGCC CAGTGAGCGT CTGCAGCTTC TCTCGGGGAC 1240

AAGCTTCCCC AGGAAGGCTG AGAGGCAGCT GCATCTGCTC CAGATGTTCT 1290

GCTTTCACCT AAAAGGCCCT GGGGAAGGGA TACACAGCAC TGGAGATTGT 1340
AAAATTTAG GAGCTATTT TTTTAACCT ATCAGCAATA TTCATCAGAG 1390
5 CAGCTAGCGA TCTTGGTCT ATTTCGGT AAAATTTGAA AATCACTAAT 1440
TCT 1443

10 (2) INFORMATION FOR SEQ ID NO:9:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 352 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Glu Leu Thr Asp Leu Leu Leu Ala Ala Met Leu Leu Ala Val
-21 -20 -15 -10

25 Ala Arg Leu Thr Leu Ser Ser Pro Val Ala Pro Ala Cys Asp Pro
-5 1 5

30 Arg Leu Leu Asn Lys Leu Leu Arg Asp Ser His Leu Leu His Ser
10 15 20

35 C
Cont Arg Leu Ser Gln Cys Pro Asp Val Asp Pro Leu Ser Ile Pro Val
25 30 35

40 Leu Leu Pro Ala Val Asp Phe Ser Leu Gly Glu Trp Lys Thr Gln
40 45 50

45 Thr Glu Gln Ser Lys Ala Gln Asp Ile Leu Gly Ala Val Ser Leu
55 60 65

50 Leu Leu Glu Gly Val Met Ala Ala Arg Gly Gln Leu Glu Pro Ser
70 75 80

55 Cys Leu Ser Ser Leu Leu Gly Gln Leu Ser Gly Gln Val Arg Leu
85 90 95

60 Leu Leu Gly Ala Leu Gln Gly Leu Leu Gly Thr Gln Gly Arg Thr
100 105 110

65 45 Thr Ala His Lys Asp Pro Asn Ala Leu Phe Leu Ser Leu Gln Gln
115 120 125

70 50 Leu Leu Arg Gly Lys Val Arg Phe Leu Leu Leu Val Glu Gly Pro
130 135 140

75 45 Thr Leu Cys Val Arg Arg Thr Leu Pro Thr Thr Ala Val Pro Ser
145 150 155

Ser Thr Ser Gln Leu Leu Thr Leu Asn Lys Phe Pro Asn Arg Thr
 160 165 170
 Ser Gly Leu Leu Glu Thr Asn Phe Ser Val Thr Ala Arg Thr Ala
 5 175 180 185
 Gly Pro Gly Leu Leu Ser Arg Leu Gln Gly Phe Arg Val Lys Ile
 190 195 200
 Thr Pro Gly Gln Leu Asn Gln Thr Ser Arg Ser Pro Val Gln Ile
 10 205 210 215
 Ser Gly Tyr Leu Asn Arg Thr His Gly Pro Val Asn Gly Thr His
 220 225 230
 15 Gly Leu Phe Ala Gly Thr Ser Leu Gln Thr Leu Glu Ala Ser Asp
 235 240 245
 Ile Ser Pro Gly Ala Phe Asn Lys Gly Ser Leu Ala Phe Asn Leu
 20 250 255 260
 Gln Gly Gly Leu Pro Pro Ser Pro Ser Leu Ala Pro Asp Gly His
 265 270 275
 25 Thr Pro Phe Pro Pro Ser Pro Ala Leu Pro Thr Thr His Gly Ser
 280 285 290
 Pro Pro Gln Leu His Pro Leu Phe Pro Asp Pro Ser Thr Thr Met
 30 295 300 305
 Pro Asn Ser Thr Ala Pro His Pro Val Thr Met Tyr Pro His Pro
 310 315 320
 Arg Asn Leu Ser Gln Glu Thr
 35 325 330 331

(2) INFORMATION FOR SEQ ID NO:10:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 331 amino acids
 (B) TYPE: Amino Acid
 (D) TOPOLOGY: Linear

45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Ser Pro Val Ala Pro Ala Cys Asp Pro Arg Leu Leu Asn Lys Leu
 1 5 10 15
 Leu Arg Asp Ser His Leu Leu His Ser Arg Leu Ser Gln Cys Pro
 50 20 25 30

Asp Val Asp Pro Leu Ser Ile Pro Val Leu Leu Pro Ala Val Asp
 35 40 45

Phe Ser Leu Gly Glu Trp Lys Thr Gln Thr Glu Gln Ser Lys Ala
 5 50 55 60

Gln Asp Ile Leu Gly Ala Val Ser Leu Leu Leu Glu Gly Val Met
 65 70 75

10 Ala Ala Arg Gly Gln Leu Glu Pro Ser Cys Leu Ser Ser Leu Leu
 80 85 90

Gly Gln Leu Ser Gly Gln Val Arg Leu Leu Leu Gly Ala Leu Gln
 15 95 100 105

Gly Leu Leu Gly Thr Gln Gly Arg Thr Thr Ala His Lys Asp Pro
 110 115 120

Asn Ala Leu Phe Leu Ser Leu Gln Gln Leu Leu Arg Gly Lys Val
 20 125 130 135

Arg Phe Leu Leu Leu Val Glu Gly Pro Thr Leu Cys Val Arg Arg
 140 145 150

25 Thr Leu Pro Thr Thr Ala Val Pro Ser Ser Thr Ser Gln Leu Leu
 155 160 165

Thr Leu Asn Lys Phe Pro Asn Arg Thr Ser Gly Leu Leu Glu Thr
 30 170 175 180

Asn Phe Ser Val Thr Ala Arg Thr Ala Gly Pro Gly Leu Leu Ser
 185 190 195

Arg Leu Gln Gly Phe Arg Val Lys Ile Thr Pro Gly Gln Leu Asn
 35 200 205 210

Gln Thr Ser Arg Ser Pro Val Gln Ile Ser Gly Tyr Leu Asn Arg
 215 220 225

40 Thr His Gly Pro Val Asn Gly Thr His Gly Leu Phe Ala Gly Thr
 230 235 240

Ser Leu Gln Thr Leu Glu Ala Ser Asp Ile Ser Pro Gly Ala Phe
 45 245 250 255

Asn Lys Gly Ser Leu Ala Phe Asn Leu Gln Gly Gly Leu Pro Pro
 260 265 270

Ser Pro Ser Leu Ala Pro Asp Gly His Thr Pro Phe Pro Pro Ser
 50 275 280 285

Pro Ala Leu Pro Thr Thr His Gly Ser Pro Pro Gln Leu His Pro
290 295 300

5 Leu Phe Pro Asp Pro Ser Thr Thr Met Pro Asn Ser Thr Ala Pro
305 310 315

His Pro Val Thr Met Tyr Pro His Pro Arg Asn Leu Ser Gln Glu
320 325 330

10 Thr
331

(2) INFORMATION FOR SEQ ID NO:11:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Ser Pro Ala Pro Pro Ala Cys Asp Pro Arg Leu Leu Asn Lys Leu
1 5 10 15

25 Leu Arg Asp Asp Xaa Val Leu His Gly Arg Leu
20 25 26

C
cont. (2) INFORMATION FOR SEQ ID NO:12:

30 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Ser Pro Ala Pro Pro Ala Xaa Asp Pro Arg Leu Leu Asn Lys Leu
1 5 10 15

40 Leu Arg Asp Asp His Val Leu His Gly Arg
20 25

(2) INFORMATION FOR SEQ ID NO:13:

45 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Xaa Pro Ala Pro Pro Ala Xaa Asp Pro Arg Leu Xaa Asn Lys
1 5 10 14

5 (2) INFORMATION FOR SEQ ID NO:14:

10 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Pro Arg Leu Leu Asn Lys Leu Leu Arg
1 5 9

15 (2) INFORMATION FOR SEQ ID NO:15:

20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 45 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

1 C
Cont. 1 GCCGTGAAGG ACGTGGTCGT CACGAAGCAG TTTATTTAGG AGTCG 45

30 (2) INFORMATION FOR SEQ ID NO:16:

35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CCNGCNCCNC CNGCNTGYGA 20

45 (2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

NCCRTGNARN ACRTGRTCRT C 21

(2) INFORMATION FOR SEQ ID NO:18:

5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 69 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

15 CCAGCGCCGC CAGCCTGTGA CCCCGACTC CTAAATAAAC TGCTCGTGA 50
TGACCACGTT CAGCACGGC 69

(2) INFORMATION FOR SEQ ID NO:19:

20 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 69 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CCAGCACCTC CGGCATGTGA CCCCGACTC CTAAATAAAC TGCTTCGTGA 50
30 CGACCACGTC CATCACGGC 69

(2) INFORMATION FOR SEQ ID NO:20:

35 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 69 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

45 CCAGCACCGC CGGCATGTGA CCCCGACTC CTAAATAAAC TGCTTCGTGA 50
CGATCATGTC TATCACGGT 69

(2) INFORMATION FOR SEQ ID NO:21:

50 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: Nucleic Acid

5 (C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

15 GCTAGCTCTA GAAATTGCTC CTCGTGGTCA TGCTTCT 37

20 (2) INFORMATION FOR SEQ ID NO:22:

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
30 (D) TOPOLOGY: Linear

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

40 CAGTCTGCCG TGAAGGACAT GG 22

45 (2) INFORMATION FOR SEQ ID NO:23:

50 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

60 Ser Pro Ala Pro Pro Ala Cys Asp Leu Arg Val Leu Ser Lys Leu
1 5 10 15

65 Leu Arg Asp Ser His Val Leu His
20 23

70 (2) INFORMATION FOR SEQ ID NO:24:

75 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

80 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

85 Ser Arg Leu Ser Gln Cys Pro Glu Val His Pro Leu Pro Thr Pro
1 5 10 15

90 Val Leu Leu Pro Ala Val Asp Phe
20 23

95 (2) INFORMATION FOR SEQ ID NO:25:

5
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

10
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Ser Leu Gly Glu Trp Lys Thr Gln Met Glu Glu Thr Lys Ala Gln
1 5 10 15
Asp Ile Leu Gly Ala Val Thr Leu
20 23

15
(2) INFORMATION FOR SEQ ID NO:26:

20
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

25
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Leu Leu Glu Gly Val Met Ala Ala Arg Gly Gln Leu Gly Pro Thr
1 5 10 15
Cys Leu Ser Ser Leu
20

30
(2) INFORMATION FOR SEQ ID NO:27:

35
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

40
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Leu Gly Gln Leu Ser Gly Gln Val Arg Leu Leu Leu Gly Ala Leu
1 5 10 15
Gln Ser Leu
18

45
(2) INFORMATION FOR SEQ ID NO:28:

50
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Leu Gly Thr Gln Leu Pro Pro Gln Gly Arg Thr Thr Ala His Lys
1 5 10 15

5 Asp Pro Asn Ala Ile Phe
20 21

^ (2) INFORMATION FOR SEQ ID NO:29:

10 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Leu Ser Phe Gln His Leu Leu Arg Gly Lys Val Arg Phe Leu Met
1 5 10 15

20 Leu Val Gly Gly Ser Thr Leu Cys Val Arg
20 25

(2) INFORMATION FOR SEQ ID NO:30:

25 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Ser Pro Ala Pro Pro Ala Cys Asp Pro Arg Leu Leu Asn Lys Leu
1 5 10 15

35 Leu Arg Asp Asp His Val Leu His Gly Arg
20 25